International Journal of Computer Assisted Radiology and Surgery (2019) 14:249–257
https://doi.org/10.1007/s11548-018-1876-6

ORIGINAL ARTICLE

Classification of contrast-enhanced spectral mammography (CESM) images

Shaked Perek[1] · Nahum Kiryati[1] · Gali Zimmerman-Moreno[2] · Miri Sklair-Levy[2] · Eli Konen[2] · Arnaldo Mayer[2]

Received: 6 June 2018 / Accepted: 13 October 2018 / Published online: 26 October 2018

Abstract

Purpose Contrast-enhanced spectral mammography (CESM) is a recently developed breast imaging technique. CESM relies on dual-energy acquisition following contrast agent injection to improve mammography sensitivity. CESM is comparable to contrast-enhanced MRI in terms of sensitivity, at a fraction of the cost. However, since lesion variability is large, even with the improved visibility provided by CESM, differentiation between benign and malignant enhancement is not accurate and a biopsy is usually performed for final assessment. Breast biopsies can be stressful to the patient and are expensive to healthcare systems. Moreover, as the biopsies results are most of the time benign, a specificity improvement in the radiologist diagnosis is required. This work presents a deep learning-based decision support system, which aims at improving the specificity of breast cancer diagnosis by CESM without affecting sensitivity.

Methods We compare two analysis approaches, fine-tuning a pretrained network and fully training a convolutional neural network, for classification of CESM breast mass as benign or malignant. Breast Imaging Reporting and Data Systems (BIRADS) is a radiological lexicon, used with breast images, to categorize lesions. We improve each classification network by incorporating BIRADS textual features as an additional input to the network. We evaluate two ways of BIRADS fusion as network input: feature fusion and decision fusion. This leads to multimodal network architectures. At classification, we also exploit information from apparently normal breast tissue in the CESM of the considered patient, leading to a patient-specific classification.

Results We evaluate performance using fivefold cross-validation, on 129 randomly selected breast lesions annotated by an experienced radiologist. Each annotation includes a contour of the mass in the image, biopsy-proven label of benign or malignant lesion and BIRADS descriptors. At 100% sensitivity, specificity of 66% was achieved using a multimodal network, which combines inputs at feature level and patient-specific classification.

Conclusions The presented multimodal network may significantly reduce benign biopsies, without compromising sensitivity.

Keywords Deep learning · Contrast-enhanced spectral mammography (CESM) · Breast cancer · Multimodal neural networks · Computer vision

Introduction

Imaging-based breast screening has played a major role in decreasing breast cancer mortality. Mammography is the only FDA-approved breast cancer screening method and is widely used for this purpose. However, it suffers from poor lesion visibility in dense breasts [1,2] (Fig. 1, left).

Contrast-enhanced spectral mammography (CESM) provides anatomical and functional imaging of breast tissue. A contrast agent is administered intravenously about two minutes before a set of low- and high-energy images is acquired sequentially, while the breast remains fixed and compressed [3]. The two images are combined to enhance areas of contrast uptake while canceling anatomic noise. The resulting lesion visibility (Fig. 1, center) is comparable to the one achieved in contrast-enhanced MRI (Fig. 1, right) but at a fraction of the cost [4,5]. The use of CESM was shown to improve lesion visibility in dense breasts [3], thus raising the diagnostic accuracy as compared to conventional mammography [6]. However, in many cases, sufficient confi-

✉ Shaked Perek
shaked.rose@gmail.com

[1] School of Electrical Engineering, Tel Aviv University, Tel Aviv, Israel

[2] Sheba Medical Center, Ramat Gan, Israel

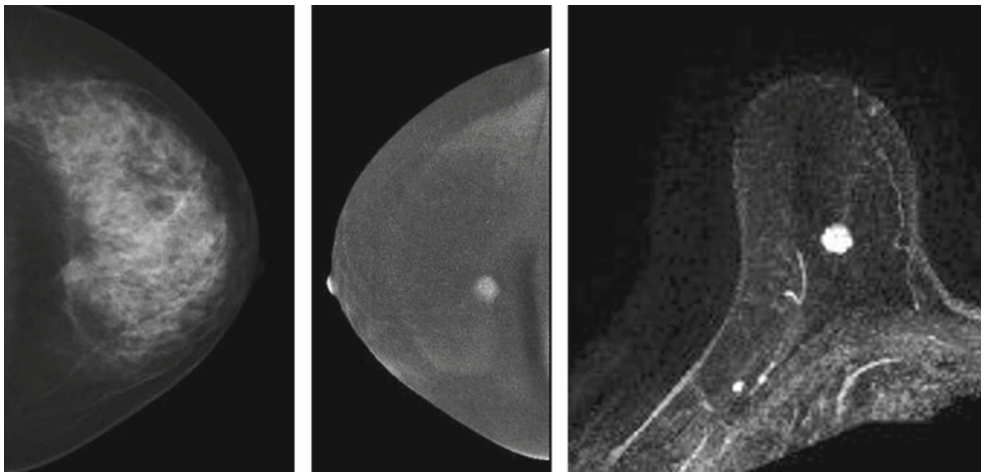

Fig. 1 Lesion visibility in dense breast: (left) classical mammography; (center) CESM; (right) contrast-enhanced MRI

dence to differentiate between benign and malignant CESM lesions cannot be reached by the radiologist and the patient is referred for biopsy. Besides being stressful for the patient, the biopsy adds cost to the diagnostic process. More than half of the breast biopsies are benign, suggesting a lot of room for specificity improvement.

Computer-assisted diagnosis (CADx) has been used for many years with mammography. Brake et al. [7] used contrast, intensity and location for classification of suspicious regions. Other image characteristics, such as morphology and texture features, were employed to aid the decision [8,9]. The improvement achieved by CAD in mammography was still insufficient, mainly due to the dense breast issue [10].

In the field of CESM, automatic image analysis systems are rare, as the imaging technique is still relatively young. Prediction of the malignant/benign label for an unseen CESM image is a challenging task. Breast cancer has many different forms, and a large dataset would be required to account for all the different variations (Fig. 2). This may prove problematic for a relatively young modality.

Mateos et al. [11] used temporal contrast-enhanced mammography (TCEM) to investigate the utility of texture features to lesion diagnosis. Textural descriptors from the Gray Level Co-occurrence Matrix (GLCM) were examined, and correlation was shown between malignancy and a combination of entropy and diagonal momentum. Cheung et al. [12] showed that CESM can assist in classifying malignant microcalcification without an associated mass, determining the size of the cancer and enabling more accurate biopsy location in the case of multiple microcalcifications. A recent study [13] presents the benefits of CAD-CESM for improving radiologists' diagnosis performance. The authors use lesion ROIs and extract 236 textural features, using image intensity, Gabor filters, GLCM and more. Features are then refined with sequential forward feature selection (SFFS) [14], and most significant features are classified using a support vector machine (SVM) [15]. These findings support our hypothesis that there is much useful quantitative information that can be extracted from contrast-enhanced mammography and leveraged toward increasing the diagnostic performance.

Breast Imaging Reporting and Data Systems (BIRADS) [16] is a radiologic lexicon that provides assessment and description for suspicious findings in the breast. BIRADS enable better communication between radiologists by providing clear definitions for lesion morphology and a scoring system for severity that aid in predicting malignancy likelihood. A specific BIRADS lexicon is fitted to each imaging technique (mammography, ultrasound and MRI). For CESM images, it was shown that MRI BIRADS descriptors are applicable [17].

In this paper, we propose and validate multimodal neural classification methods that seamlessly combine CESM pixel information with textual BIRADS descriptors provided by the radiologist.

By enhancing the information used for lesion classification (text and pixels), we aim at increasing the diagnostic specificity, thereby reducing benign biopsies without compromising sensitivity. The advantages of multimodal networks have been previously proven on multiple tasks, ranging from video classification [18], fusing video and audio, hand gesture recognition [19] and image retrieval [20], combining medical images with textual descriptions of the image. In these works, using features from different modalities improves classification results. Different modalities complement each other and contribute to precise prediction.

The proposed multimodal classification methods are presented in "Methods" section. Data acquisition, experiments and results are presented in "Experiments" section. Conclusions are summarized in "Conclusion" section.

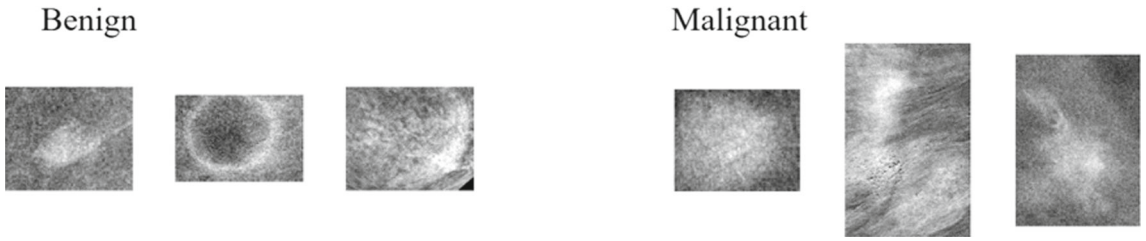

Fig. 2 Crops for sample lesions in the dataset. Breast lesions morphology and size are diverse; hence, CESM dataset has a large variability. Benign lesions (left), malignant lesions (right)

Methods

The networks

We compare two CNN methods, fine-tuning pretrained AlexNet [21] and *end to end* training of a deep network with interchanging layers of convolution and max pooling.

CESM images have high resolution, typically about 3000×2000 pixels. Such large images, if used as network input, lead to high computational cost and require a large number of parameters. A common approach is to resize the images as a preprocessing step. However, lesion texture and pattern are subtle and are likely to be harmed by the resizing operation. Therefore, it is more convenient to use a set of sample image patches, instead of the whole mammogram, as input to the network.

The dataset includes lesions of different shapes and sizes. Therefore, the number of sample patches used has to account for different cases. If a lesion is small, a large number of patches will result in overlapping patches that contain similar information. Conversely, a large lesion area with a small number of patches will not sufficiently cover the different parts of the lesion. For each grayscale image, we randomly select 200 pixels belonging to the lesion mask and extract a square patch (227×227 pixels) around each sampled pixel. This method satisfies the need for a large dataset, required for deep learning approaches, and preserves the original image resolution within each patch.

At testing time, inference is done on all lesion patches and final classification score for a considered lesion is the mean score computed over its 200 sample patches.

In order to classify each lesion, the problem is formulated as an optimization task that minimizes the cross-entropy cost function defined by:

$$\text{Loss} = -\sum_{i=1}^{C} L_i \log(y_i) \tag{1}$$

where $L_i \in [0, 1]$ is the label, $C = 2$ number of classes and y_i is the softmax classification score on network output $Z_{i=\text{malignant/benign}}$

$$y_i = \frac{e^{Z_i}}{\sum_{d=1}^{C} e^{Z_d}} \tag{2}$$

Fine-tuned AlexNet

The practice of fine-tuning a pretrained network is common, usually an extension to transfer learning [22,23]. Previous work has shown that despite the difference between natural and medical images, fine-tuning a pretrained network outperforms networks trained from scratch, due to lack of large-scale labeled medical datasets [22].

Alexnet [21] was trained on the ImageNet dataset, a mass collection of about 14 million natural images with 1000 image classes. It consists of five convolution layers with ReLU activation function, three pooling layers and three fully connected layers with dropout. In this work, in order to adapt the network to the classification of breast lesions, we change the last fully connected layer to output only two classes and fine-tune the last two layers of AlexNet. We will refer to this method as FT-AlexNet, and its architecture is shown in Fig. 3a (dashed arrow not part of it).

End to end training

To explore an alternative to FT-AlexNet, we define a new CNN, hereon called RawNet. It consists of five convolution layers with batch normalization (BN) [24] after each convolution and ReLU activation, three max-pooling layers and two fully connected layers. This network has around 240 K parameters. The network diagram can be seen in Fig. 3b.

Patient-specific classification

In order to further improve lesion classification performance, we propose a patient-specific fine-tuning step. For this purpose, we assume that breast tissue located outside of the contoured lesions can be considered as normal, non-malignant tissue. This makes sense as the radiologist is requested to contour all the suspected lesions in the CESM mammogram. At testing, for each patient, the assumed non-malignant tissue is sampled with square patches and labeled as benign (see Fig. 4). These samples are considered *normal*

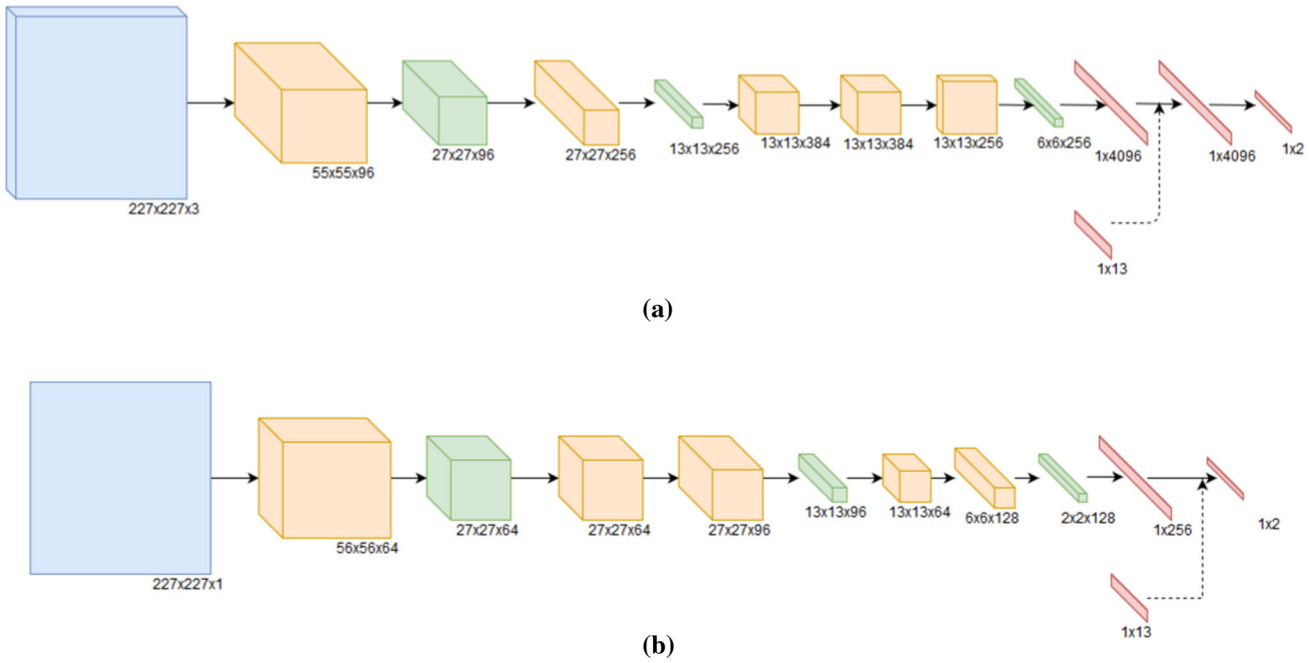

Fig. 3 CNN architectures with each layer's output size. **a** FT-AlexNet. **b** RawNet. Input grayscale image (blue), convolution layers (orange), max pooling layers (green), fully connected (red). The dashed line is concatenation of BIRADS feature vector to the output of the first fully connected layer, as used in the feature fusion approach

patches. The network that was already trained on the whole training set is further fine-tuned for an additional epoch using the patient-specific non-malignant tissues, before the lesion patches are classified. This is a personal classification step applied per patient.

Multimodal classification

BIRADS encoding

The BIRADS are provided by the radiologist as textual descriptors and consequently converted into numerical inputs usable by the neural network as explained below. In this work, four categories are used from the CESM BIRADS lexicon (Table 1): (1) mass shape, providing a basic description of the lesion form. (2) mass margin, defining qualitatively the lesion boundary. For example, spiculated margins are an indication of malignancy. (3) Mass internal enhancement, qualitatively describing the spatial distribution of hyper-intensity inside the lesion. For example, a cyst, which is in most cases benign, will be bright only on the margin (rim enhancement). (4) Background Parenchymal Enhancement (BPE), quantifying the enhancement of breast fibroglandular tissue outside the lesion. The BIRADS textual descriptors are converted into one-hot vectors and, after concatenation, result in a 13 elements feature vector. Details for each BIRAD descriptor are given in Table 1, the last columns showing its one-hot vector conversion.

BIRADS fusion

We examine two ways of incorporating the BIRADS vector to obtain multimodal classification. In the first approach, denoted feature fusion, the BIRADS binary vector and the output of a fully connected layer are concatenated before entering the next layer of the network. In both networks, FT-AlexNet and RawNet, concatenation is performed with the output of the first fully connected layer, as illustrated by the dashed arrow in Fig. 3a, b.

In the second approach, denoted decision fusion, the classification scores from two classifiers, one per modality, are merged to provide the final benign/malignant decision. For this purpose, the aforementioned CNN (without BIRADS) and an SVM classifier, denoted BSVM, trained only on BIRADS features are implemented to provide, each, a separate classification score [15]. The final lesion malignancy classification can merge the classification scores in several ways:

Max fusion Given a lesion to be classified, each classifier outputs a probability vector, denoted by $[p(x_{\text{benign}}), p(x_{\text{malignant}})]$, and quantifies the probability that the lesion is benign or malignant, respectively. In order to stay on the

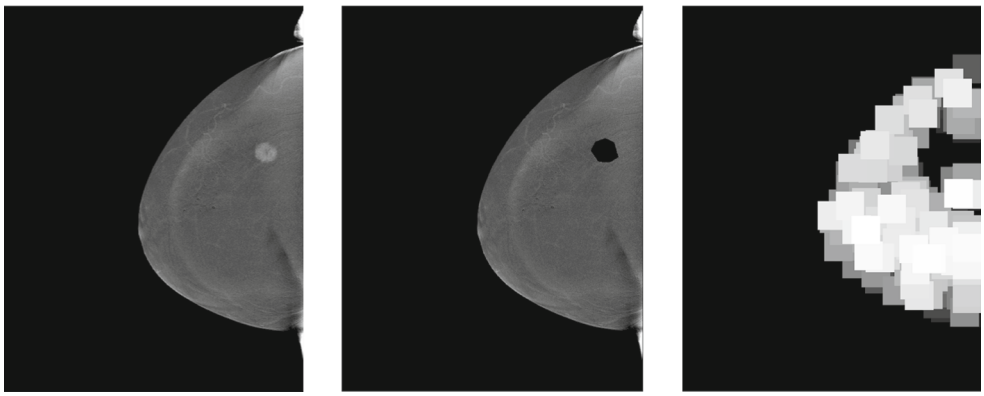

Fig. 4 Non-malignant tissue patches acquisition. (Left) CESM original image, (middle) CESM image without lesion area, (right) sampling of 200 non-malignant patches from areas external to the lesion contour

Table 1 Four BIRADS categories taken from the BIRADS atlas used as textual input to the network. Columns 4 to the end of the table are the one-hot vectors defined per each descriptor. BPE has four descriptors as opposed to the other categories which have 3

BPE	Level	Minimal	1	0	0	0
		Mild	0	1	0	0
		Moderate	0	0	1	0
		Marked	0	0	0	1
Mass	Shape	Oval	1	0	0	
		Round	0	1	0	
		Irregular	0	0	1	
	Margin	Circumscribed	1	0	0	
		Not circumscribed irregular	0	1	0	
		Not circumscribed spiculated	0	0	1	
	Internal enhancement	Homogeneous	1	0	0	
		Heterogeneous	0	1	0	
		Rim enhancement	0	0	1	

safe side, we take the maximum of the probabilities to define the fused score for the malignant class.

$$\text{score}_{\text{malignant}} = \max(p(x_{\text{malignant}}|\text{BSVM}),\ p(x_{\text{malignant}}|\text{CNN})) \tag{3}$$

where $p(x_{\text{malignant}}|\text{BSVM})$ is the malignant class probability given by BSVM and $p(x_{\text{malignant}}|\text{CNN})$ is the malignant class probability given by CNN.

Linear combination Another option is to compute a weighted average of the probabilities. The coefficient α controls the contribution of each classifier.

$$\text{score}_{\text{malignant}} = \alpha \cdot p(x_{\text{malignant}}|\text{BSVM}) + (1-\alpha) \cdot p(x_{\text{malignant}}|\text{CNN}) \tag{4}$$

SVM classifier Alternatively, an SVM classifier can be used for the scores fusion task. It defines a hyperplane maximizing the margin between the plane and the two classes it separates. The margin is defined by the distance from the hyperplane to the nearest data point. We define the new feature vector set of size Nx2 as the malignancy probabilities from each classifier. Since our dataset is small, the training set for this classifier is the same training set used for the network and BSVM classifier.

Experiments

Data and preprocessing

A CESM dataset consisting of 54 patient with 129 images of left/right cranio-caudal (CC) and mediolateral oblique (MLO) views was acquired on a GE Senographe Essential mammography. The dataset contains 73 images with benign lesions and 56 malignant lesion images. The CESM is a combination of high and low energy exposures (45–50 kVp, 26–30 kVp, respectively) acquired two minutes after contrast medium (iodine-based) injection. The typical image size is 2394×3062. Data annotation was done by an expert radiologist and includes a rough contour of the lesion and BIRADS

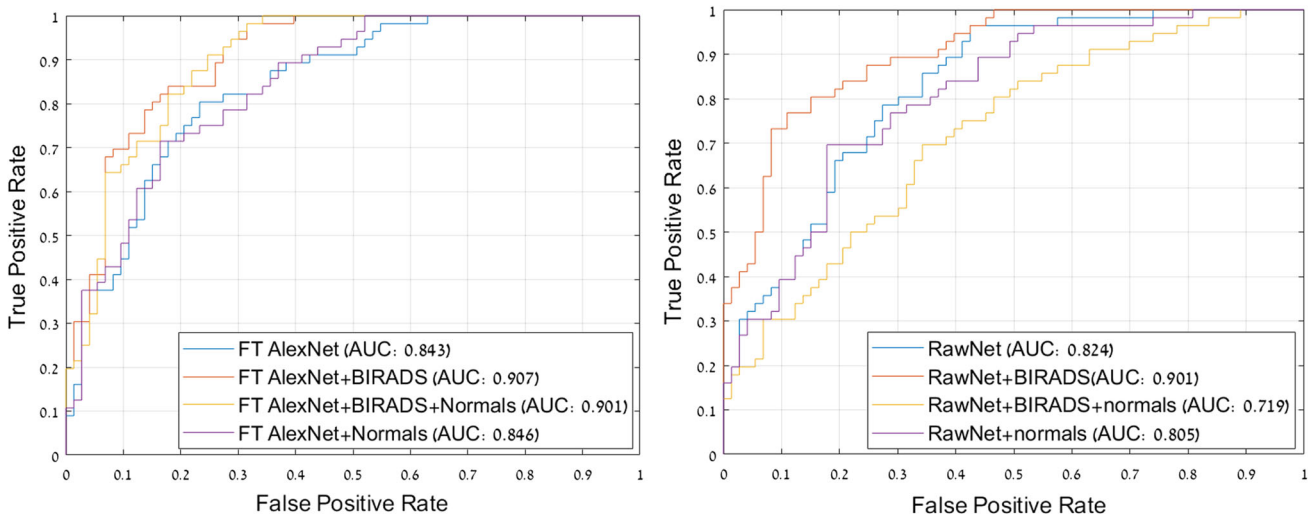

Fig. 5 Comparison between ROC of the two classification networks. (left) FT-Alexnet, (right) RawNet. Four curves are presented with their matching AUC: (1) pixels input alone (FT-Alexnet/RawNet, blue),(2) *feature* fusion of pixels and BIRADS (FT-AlexNet/RawNet + BIRADS, red), (3) *feature* fusion of pixels and BIRADS with network fine-tuning using normal patches (FT-AlexNet/RawNet + BIRADS + normal, yellow) and (4) pixels input alone with network fine-tuning using normal patches (FT-AlexNet/RawNet + normal, purple)

descriptors for each image. The lesions in the dataset were all biopsy proven, to provide the ground truth (GT) labels as malignant/benign.

CESM images are uint16 numeric type while the input to AlexNet [21] consists of three uint8 channels. In order to make CESM images readable by the AlexNet network, we have to map the dynamic range of the CESM into the numerical format of a single network channel. A compression of the CESM dynamic range is therefore necessary . To minimize compression and potential contrast loss, the values are clipped between 1900 and 2500 to contain most of the histogram energy and only then rescaled between 0 and 255. Eventually, the mean value is subtracted from each image patch before being duplicated and fed to the three input channels of FT-AlexNet simultaneously.

A common technique to increase the size of the dataset is to use augmentations. As addition to the various translations created by the patch-based method, each patch in the training set was also rotated by multiples of 90° and flipped up–down or left–right to represent a variety of possible lesion positions.

The parameters

The hyperparameters used in all experiments are: the Adam optimizer [25] for network training, with a batch size of 200, learning rate of 0.0001 and dropout 0.5. For the linear combination option in the decision fusion method, described in "BIRADS fusion" section, we set α to be [0.5, 0.3]. The first option (denoted *Mean Malignant*) is an average between BIRADS' and network's decision. The second option (denoted *Linear Combination Malignant*), 0.3, gives more weight to the network decision which is based on the image input. All the experiments are implemented using five-fold cross-validation.

Validation schemes

In a first experiment (Fig. 5), we compare the receiver operating characteristics (ROC) and corresponding area under the curve (AUC) for FT-AlexNet (left) and RawNet (right) networks in four different settings: (1) pixels input alone (FT-Alexnet/RawNet,blue), (2) *feature* fusion of pixels and BIRADS (FT-AlexNet/RawNet + BIRADS, red), (3) *feature* fusion of pixels and BIRADS with network fine-tuning using normal patches (FT-AlexNet/RawNet+BIRADS + normal, yellow) and (4) pixels input alone with network fine-tuning using normal patches (FT-AlexNet/RawNet + normal, purple).

Among both the considered architectures, FT-AlexNet+ BIRADS (red curve) achieves the highest AUC, at 0.9. For pixel data only (no BIRADS), at the maximum sensitivity working point (no false negative), FT-AlexNet (blue) achieves the best specificity at 0.37, with only 0.27 for RawNet (Table 2). This means that benign biopsies can be reduced by up to 37% using pixel-only FT-AlexNet. By adding textual BIRADS information, we see that multimodal FT-AlexNet+BIRADS can do significantly better and reduces by up to 60% the benign biopsies count (Table 2). Further improvement is obtained by network fine-tuning using normal patches (see "Patient-specific classification" section), as the potential reduction in benign biopsies reaches 66% (Table 2). Comparatively, the addition of BIRADS

Table 2 Specificity results summary for the sensitivity=100% working point

Network	Feature fusion				Decision fusion							
	Pixels-only		Multimodal		SVM		Max		Linear comb. $\alpha = 0.5$		Linear comb. $\alpha = 0.3$	
FT-AlexNet	37	48	60	**66**	50	51	45	47	58	58	51	52
RawNet	27	10	53	20	**56**	**56**	44	47	41	45	36	36

For each experiment we present two results (separated by a vertical line): *without* and *with* normal patches fine-tuning (left and right, respectively). Best result are in bold for each network

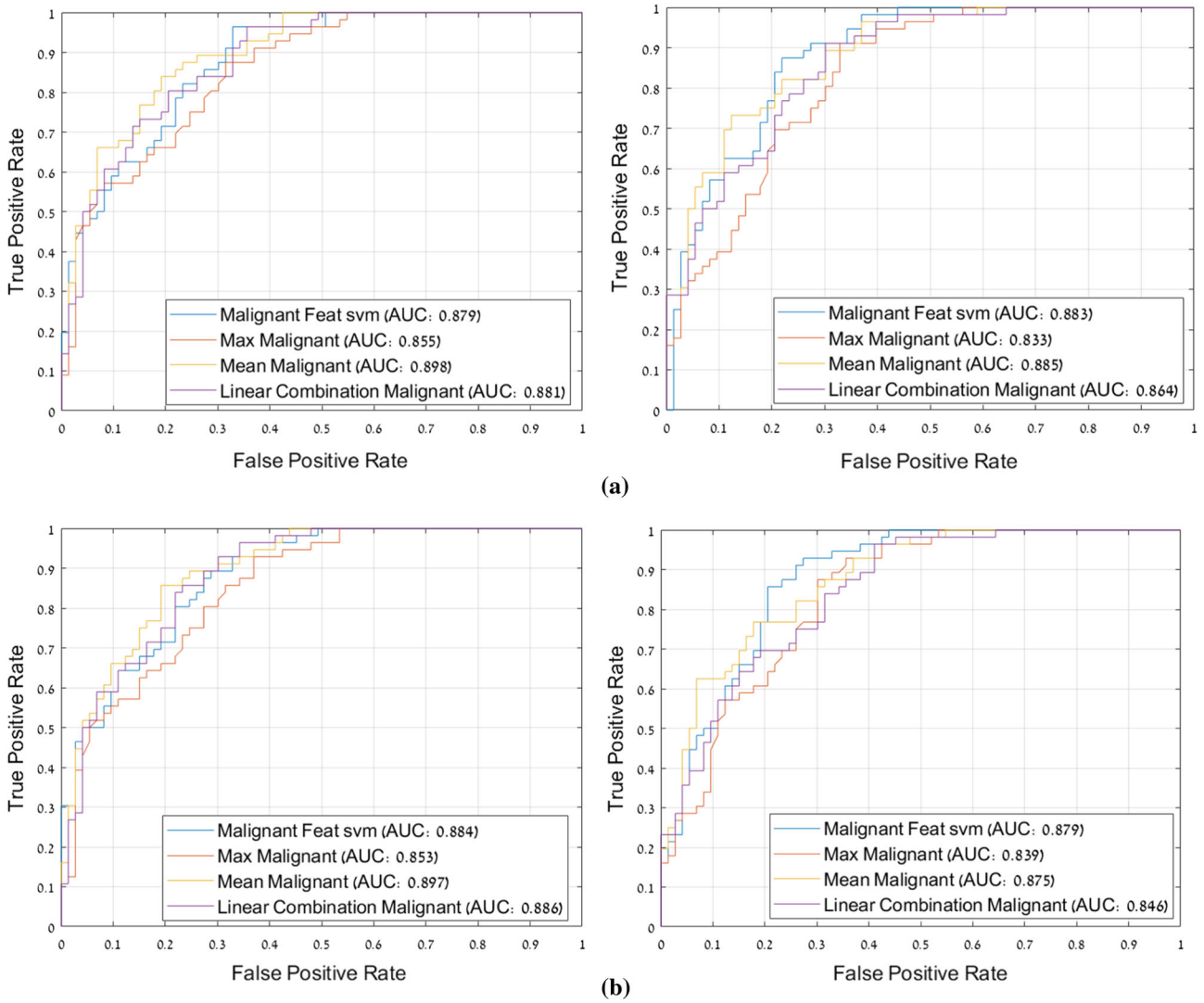

Fig. 6 Comparison between classification ROC of the proposed multimodal networks with decision fusion. (left) FT-Alexnet, (right) RawNet. **a** *Without* normal patches fine-tuning; **b** *with* normal patches fine-tuning. ROCs are given in four decision fusion alternatives: SVM (malignant feat SVM, blue), max fusion (max malignant, red), linear combination with $\alpha = 0.5$ (mean malignant, yellow) and linear combination with $\alpha = 0.3$ (linear combination malignant, purple). For RawNet, the highest AUC and best SPC score are achieved by SVM fusion. For FT-AlexNet, the best score is achieved by mean malignant

information to RawNet yields a smaller potential in benign biopsies reduction, at 53% (Table 2). Moreover, further adding the normal patches fine-tuning results in a decline of both the AUC (0.71 for RawNet+BIRADS+normal, 0.8 for RawNet+normal) and the specificity at maximum sensitivity working point (10% and 20%, respectively, in Table 2).

In a second experiment (Fig. 6), the *decision* fusion approach is evaluated as an alternative to the *feature* fusion of the first experiment. Fig. 6a, b shows *decision* fusion results for FT-AlexNet (left) and RawNet (right) *without* and *with* normal patches fine-tuning, respectively. For FT-AlexNet (left), the best decision fusion results are achieved by using a *linear combination* wherein $\alpha = 0.5$ (denoted Mean Malignant, yellow) . The AUC *with* (b) or *without* (a) normal patches fine-tuning is 0.89, with up to 58% (Table 2) reduction in benign biopsies.

For RawNet (right), the best *decision* fusion results (with and without normal patches) are achieved for an SVM classifier (denoted Malignant Feat SVM, blue): AUC = 0.88 and up to 56% (Table 2) reduction in benign biopsies.

In both fusion methods, the improvement in classification with multimodal networks was significant. This shows the benefits of the BIRADS addition. Even though the baseline FT-AlexNet is fairly capable at malignancy prediction, BIRADS information is an important addition. Patient-specific classification has proved to be important as well, but only with the more robust FT-AlexNet. Addition of normal patches information is beneficial in the feature fusion method, which achieved the highest specificity score and is most suitable for benign biopsy reduction.

Conclusions

CESM is an effective and economical modality for breast cancer diagnosis. A suspicious lesion in a CESM image usually leads to biopsy, to determine whether it is benign or malignant. Biopsies are invasive and add cost to the diagnostic process. The purpose of this work is to increase the specificity of the CESM diagnostic so that benign biopsies can be significantly reduced, without compromising sensitivity.

The task of lesion classification in CESM images was scarcely addressed by the research community. This is due to the lack of a large, biopsy-proven database of suspicious lesion images obtained with the new CESM modality.

The significant variation in visual CESM characteristics of both benign and malignant lesions makes classification difficult. We overcome these challenges using a deep learning approach to handle the complexity of the available CESM data.

We explored ways of fusing textual information given by an expert radiologist with CESM suspicious lesion images. We present a suitable multimodal network and show its diagnostic advantage. Our trained model can potentially lead to substantial reduction in benign biopsies. We show that feature fusion is preferable to decision fusion, since it enables the network to learn suitable weights per parameter. Patient-specific classification (by normal patches fine-tuning) is helpful in final assessment and achieves better malignancy prediction scores. These scores can serve as a second-reader input assisting medical experts in their decisions.

In this work, FT-AlexNet was essential to achieving the best results for potential benign biopsy reduction. A fully trained network could not meet the requirements in a sufficient manner, since large CESM datasets are not yet available.

Future work will focus on gathering more data. This will be helpful in better training and more robust evaluation. It can be achieved using new methods of synthesizing data. Other pretrained networks can also be explored. In our work, we use pretrained AlexNet [21]; however, the improvements suggested here are impartial to the base network used and can be added with ease. Further research would also examine how information regarding a patient's medical history can improve classification.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent For this type of study formal consent is not required.

References

1. Carney PA, Miglioretti DL, Yankaskas BC, Kerlikowske K, Rosenberg R, Rutter CM, Geller BM, Abraham LA, Taplin SH, Dignan M, Cutter G (2003) Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. Ann Intern Med 138(3):168–75
2. Drukteinis JS, Mooney BP, Flowers CI, Gatenby RA (2013) Beyond mammography: new frontiers in breast cancer screening. Am J Med 126(6):472–479
3. Lobbes MBI, Smidt ML, Houwers J, Tjan-Heijnen VC, Wildberger JE (2013) Contrast enhanced mammography: techniques, current results, and potential indications. Clin Radiol 68(9):935–944
4. Li L, Roth R, Germaine P, Ren S, Lee M, Hunter K, Tinney E, Liao L (2017) Contrast-enhanced spectral mammography (CESM) versus breast magnetic resonance imaging (MRI): a retrospective comparison in 66 breast lesions. Diagn Interv Imaging 98:113123
5. Fallenberg EM, Dromain C, Diekmann F, Engelken F, Krohn M, Singh JM, Ingold-Heppner B, Winzer KJ, Bick U, Renz AD (2014) Contrast-enhanced spectral mammography versus MRI:

initial results in the detection of breast cancer and assessment of tumour size. Eur Radiol 24(1):256–264
6. Luczyska E, Heinze-Paluchowska S, Dyczek S, Blecharz P, Rys J, Reinfuss M (2014) Contrast-enhanced spectral mammography: comparison with conventional mammography and histopathology in 152 women. Korean J Radiol 15(6):689–696
7. Brake GM, Karssemeijer N, Hendriks JH (2000) An automatic method to discriminate malignant masses from normal tissue in digital mammograms1. Phys Med Biol 45(10):2843
8. Sampat MP, Markey MK, Bovik AC (2005) Computer-aided detection and diagnosis in mammography. Handb Image Video Process 2(1):1195–1217
9. Vasantha M, Bharathi DVS, Dhamodharan R (2010) Medical image feature, extraction, selection and classification. IJEST 1(2):2071–2076
10. Jalalian A, Mashohor SB, Mahmud HR, Saripan MIB, Ramli ARB, Karasfi B (2013) Computer-aided detection/diagnosis of breast cancer in mammography and ultrasound: a review. Clin Imaging 37(3):420–426
11. Mateos MJ, Gastelum A, Mrquez J, Brandan ME (2016) Texture analysis of contrast-enhanced digital mammography (CEDM) images. International Workshop on Digital Mammography. LNCS 9699, Springer, pp 585–592
12. Cheung YC, Tsai HP, Lo YF, Ueng SH, Huang PC, Chen SC (2016) Clinical utility of dual-energy contrast-enhanced spectral mammography for breast microcalcifications without associated mass: a preliminary analysis. Eur Radiol 26(4):1082–1089
13. Patel BK, Ranjbar S, Wu T, Pockaj BA, Li J, Zhang N, Lobbes M, Zhang B, Mitchell JR (2018) Computer-aided diagnosis of contrast-enhanced spectral mammography: a feasibility study. Eur J Radiol 98:207–213
14. Pudil P, Novoviov J, Kittler J (1994) Floating search methods in feature selection. Pattern Recognit Lett 15(11):1119–1125
15. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297
16. American College of Radiology, ACR BI-RADS Atlas 5th Edition, 125–143 (2013)
17. Knogler T, Homolka P, Hoernig M, Leithner R, Langs G, Waitzbauer M, Pinker-Domenig K, Leitner S, Helbich TH (2017) Application of BI-RADS descriptors in contrast-enhanced dual-energy mammography: comparison with MRI. Breast Care 12:212–216
18. Ngiam J, Khosla A, Kim M, Nam J, Lee H, Ng AY (2011) Multimodal deep learning. In: Proceedings of the 28th ICML
19. Wu D, Pigou L (2016) Deep dynamic neural networks for multimodal gesture segmentation and recognition. IEEE Trans Pattern Anal Mach Intell 38(8):1583–1597
20. Cao Y, Steffey S, He J, Xiao D, Tao C, Chen P, Mller H (2014) Medical image retrieval: a multimodal approach. Cancer Inf 13:CIN-S14053
21. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. NIPS 25:1097–1105
22. Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, Liang J (2016) Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE TMI 35(5):1299–1312
23. Karpathy A, Toderici G, Shetty S, Leung T, Sukthankar R, Fei-Fei L (2014) Large-scale video classification with convolutional neural networks. In: IEEE conference CVPR, pp 1725–1732
24. Ioffe S, Szegedy C (2015) Batch normalization: accelerating deep network training by reducing internal covariate shift. In: ICML, pp 448–456
25. Kingma D, Adam JB (2014) A method for stochastic optimization. arXiv preprint arXiv:1412.6980